

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/82, 15/29, A01H 5/00, C12N 5/10		A1	(11) International Publication Number: WO 95/35386 (43) International Publication Date: 28 December 1995 (28.12.95)
(21) International Application Number: PCT/GB95/01443 (22) International Filing Date: 19 June 1995 (19.06.95) (30) Priority Data: 9412123.3 17 June 1994 (17.06.94) GB 9412124.1 17 June 1994 (17.06.94) GB (71) Applicant (for all designated States except US): SCOTTISH CROP RESEARCH INSTITUTE [GB/GB]; Invergowrie, Dundee DD2 5DA (GB). (72) Inventors; and (75) Inventors/Applicants (for US only): BROWN, John, William, Slessor [GB/GB]; 49 Meadowview Drive, Inchtute, Perthshire PH14 9TQ (GB). CLARK, Gillian, Patricia [GB/GB]; 7 Lowson Road, Forfar, Angus DD8 1LB (GB). SIMPSON, Gordon, Grant [GB/GB]; Friedrich, Miescher Institute, CH-4005 Basel (CH). (74) Agent: PATTULLO, Norman; Murgitroyd & Company, 373 Scotland Street, Glasgow G5 8QA (GB).			(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

(54) Title: **EXPRESSION CONTROL POLYNUCLEOTIDES DERIVED FROM SPLICEOSOMAL PROTEIN GENE PROMOTERS**

(57) Abstract

An expression control polynucleotide capable of affecting the expression of a second polynucleotide, and derived from a spliceosomal protein gene promoter. Isolation of plant spliceosomal protein gene promoters from potato and maize is described. Partial sequences of the promoters of two potato spliceosomal protein gene promoters are disclosed.

AL

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

1 "Expression Control Polynucleotides derived from
2 Spliceosomal Protein Gene Promoters"

3
4 This invention relates to expression control
5 polynucleotides derived from spliceosomal protein gene
6 promoters.

7
8 It is already well known that organisms such as plants
9 or animals with novel characteristics can be produced
10 by introducing genes or DNA sequences from the same or
11 a related organism or by introducing genes or DNA
12 sequences from other organisms. The key to production
13 of novel phenotypes is the active expression of at
14 least a part of the introduced DNA sequence. Generally
15 expression of the introduced DNA sequence will occur
16 only in the presence of an expression control
17 polynucleotide, such as a promoter, which is compatible
18 with the host organism. Promoters are nucleic acid
19 sequences which are currently believed to regulate the
20 expression of a gene by facilitating the binding of
21 proteins required for transcription, such as RNA
22 polymerase, to a portion of the nucleic acid sequence
23 upstream of the gene. Thus for protein coding genes
24 the DNA is transcribed into mRNA (messenger RNA) which
25 is then relocated to the cytoplasm where it is

1 available for translation into polypeptide.

2

3 Whereas promoters are generally disposed upstream of
4 the genes they regulate and are thought to act by
5 providing a binding site for an RNA polymerase
6 transcription complex, another form of expression
7 control polynucleotide known as an "enhancer" sequence
8 can control the expression of a gene usually without
9 regard to its relative position or orientation.

10

11 Promoters can vary in levels of expression induction
12 and in their expression patterns. Some promoters are
13 active in a tissue-specific or developmental stage-
14 specific manner while other promoters are active
15 constantly to continually drive expression of the gene
16 they control. Such constantly active promoters are
17 called constitutive promoters. The promoters of this
18 latter class most widely used in genetic engineering
19 are the Cauliflower Mosaic Virus (CaMV) 35S RNA
20 promoter, nopaline synthetase (nos) or octopine
21 synthetase (ocs) promoters. Constitutive promoters
22 induce gene expression at a relatively constant rate.

23

24 Although the above conventional promoters generally
25 drive expression at high levels, many have the
26 disadvantage that they are derived from plant
27 infectious agents; the CaMV 35S RNA promoter is derived
28 from a plant virus and the other promoters mentioned
29 are derived from strains of *Agrobacterium*, soil-borne
30 infectious bacteria. The source of these promoters is
31 a cause of concern in transgenic plant production. In
32 addition, detailed analysis of expression patterns of
33 CaMV 35S promoter, have shown that its levels of
34 expression can vary greatly among different plant
35 tissues to the level where it is inactive in some
36 tissues and is therefore, no longer constitutive for

1 such tissue. For many biotechnological objectives,
2 constitutive expression in all cells and tissues would
3 be of great advantage.

4
5 Splicesomal proteins are believed to be present in
6 virtually all eukaryotic cells and are involved in the
7 phenomenon of pre-mRNA splicing which removes introns
8 (non-coding regions) from RNA transcripts before
9 protein production.

10
11 The present invention seeks to provide a promoter (or
12 other expression control polynucleotide) which is not
13 derived from an infectious agent and which is suitable
14 for use in the control of expression of recombinant
15 genes in the construction of transgenic organisms such
16 as plants and animals.

17
18 The present invention also seeks to provide a promoter
19 (or other expression control polynucleotide) which is
20 likely to be active throughout all or most cells of the
21 organism.

22
23 According to the present invention there is provided an
24 expression control polynucleotide at least partially
25 derived from a spliceosomal protein gene promoter.

26
27 The expression control polynucleotide of the invention
28 is capable of controlling the expression of a second
29 polynucleotide (preferably comprising a polypeptide-
30 encoding sequence) operably linked thereto.

31
32 RNA sequences which do not code for protein can also be
33 expressed eg ribozymes or anti-sense RNA.

34
35 The term "expression control polynucleotide" as used
36 herein will include promoters, enhancers or any other

1 functional equivalents or any other sequence elements
2 which affect expression of other gene sequences.

3

4 The term "polypeptide-encoding sequence" as applied
5 herein to polynucleotides means a polynucleotide
6 comprising a sequence which can be transcribed into
7 mRNA, which itself can be translated into a
8 polypeptide. The "polypeptide-encoding sequence" may
9 include non-translated portions, such as introns.

10

11 The spliceosomal protein gene promoter may be derived
12 from plants. The plants may be dicotyledonous (eg
13 potatoes), or monocotyledonous (eg maize).

14

15 The present invention also provides a recombinant
16 polynucleotide comprising an expression control
17 polynucleotide according to the invention operably
18 linked to a second polynucleotide (preferably
19 comprising a polypeptide-encoding sequence).

20

21 The present invention also provides a recombinant
22 vector containing an expression control polynucleotide
23 or a recombinant polynucleotide as defined above.

24

25 According to the present invention there is also
26 provided a method of producing a recombinant vector,
27 said method comprising ligating an expression control
28 polynucleotide into a vector or part thereof. A method
29 of producing a transformed cell by transfecting a host
30 cell using said recombinant vector forms another aspect
31 of the invention.

32

33 The present invention also provides a transformed host
34 cell containing a recombinant polynucleotide or vector
35 as defined above.

36

1 The present invention also provides a transgenic
2 organism (for example a transgenic plant) containing a
3 recombinant polynucleotide or vector as defined above.
4 The progeny (and seeds) of such transgenic organisms
5 forms a further part of the invention.

6
7 The present invention also provides a method for
8 controlling the expression of a polypeptide from a
9 nucleotide sequence encoding the polypeptide, said
10 method comprising operably linking said sequence to an
11 expression control polynucleotide of the invention.

12
13 The expression control polynucleotide of the invention
14 may comprise double- or single-stranded DNA or RNA.

15
16 Three cultures of *E.coli* (SCRI/JB/1, SCRI/JB/2 and
17 SCRI/JB/3), each containing a plasmid having an
18 expression control polynucleotide according to the
19 invention were deposited on 14 March 1994 with the
20 National Collection of Type Cultures under numbers NCTC
21 12864, NCTC 12865 and NCTC 12866 respectively.

22
23 Cultures SCRI/JB/1 and 2 contain dicotyledonous
24 spliceosomal protein gene expression control
25 polynucleotides (promoters for potato U1A and U2B"
26 genes respectively); SCRI/JB/3 contains a
27 monocotyledonous expression control polynucleotide
28 (promoter for a maize PRP8 gene).

29
30 Accordingly, the present invention also provides NCTC
31 deposits Nos 12864, 12865 and 12866 and the plasmids
32 thereof.

33
34 Recombinant DNA technology has been recognised as a
35 powerful technique not only in research but also for
36 commercial purposes. Thus, by using recombinant DNA

1 techniques (see Sambrook et al 1982 and "Principles of
2 Genetic Engineering", Old and Primrose, 5th edition,
3 1994) exogenous genetic material can be transferred to
4 a host cell and the polypeptide encoded by the
5 exogenous genetic material may be replicated by and/or
6 expressed within the host. For the purposes of
7 simplicity recombinant DNA technology is normally
8 carried out with prokaryotic micro-organisms, for
9 example bacteria such as *E. coli*, as host. However,
10 use has also been made of eukaryotic organisms, in
11 particular yeasts or algae, and in certain applications
12 eukaryotic cell cultures may also be used.

13
14 Genetic alterations to mammalian species by micro-
15 injection of genes into the pro-nuclei of single-cell
16 embryos is also well known and has been described by
17 Brinster et al, in *Cell* 27: 223-231, 1981. Thus
18 general techniques used in recombinant DNA technology
19 and the production of transgenic organisms is within
20 the scope of the skilled man.

21
22 Using a cDNA sequence (complimentary DNA - reverse
23 transcribed from mRNA), either full-length or partial,
24 as a probe for a gene of interest, the gene promoter
25 can be readily isolated by standard procedures.
26 Briefly, the cDNA probe is first used to screen a
27 genomic library to isolate a genomic clone containing
28 the promoter and coding sequence. Restriction mapping
29 and Southern blotting with the cDNA as probe delineates
30 the region of the genomic clone containing the coding
31 sequence. Sequencing of this region of the genomic
32 clone and comparison to the cDNA clone will identify
33 the translation initiation ATG codon, and if the cDNA
34 is full-length, will give an indication of the
35 transcription start site, upstream of which lies the
36 gene promoter. Important promoter sequence elements may

1 lie in excess of around 2 kbp of the transcription
2 start site. Therefore, a genomic fragment of preferably
3 1 - 5 kbp is isolated for initial testing of promoter
4 activity. For example, in the case of U1A and U2B", a
5 monoclonal antibody (mAb), 4G3, raised against a b-
6 galactosidase-human U2B" protein, was used as a probe
7 to screen a potato cDNA expression library and a full-
8 length cDNA clone was isolated (Simpson et al. 1991).
9 Screening of a potato genomic library using this cDNA
10 clone resulted in a genomic clone containing part of
11 the gene and around 15 kbp upstream sequences which
12 were used to clone the promoter in subsequent
13 experiments. The extreme similarity between U1A and
14 U2B" enabled us to isolate a full-length genomic clone
15 of the U1A gene. Since the latter clone contained
16 around 7 kbp of upstream sequences, it was possible to
17 clone the promoter from the upstream region.

18
19 The use of spliceosomal protein gene promoters to drive
20 expression of DNA sequences in transgenic plants or
21 animals has the advantage that the expression is
22 constitutive, expressed in all cell and tissue types
23 and uses naturally occurring plant or animal nucleic
24 acid sequences which are not derived from infectious
25 agents.

26
27 While further modifications and improvements may be
28 made without departing from the scope of this
29 invention, the following is a description of one or
30 more examples of the invention, with reference to the
31 accompanying drawings, in which:-

32
33 Fig. 1 shows a schematic diagram of a potato U2B"
34 genomic clone present in the cells of culture No
35 SCRI/JB/2;

36 Fig. 2 shows a schematic diagram of a potato U1A

1 genomic clone present in the cells of culture No
2 SCRI/JB/1;
3 Fig. 3 shows a schematic diagram of a maize PRP8
4 genomic clone present in the cells of culture No
5 SCRI/JB/3;
6 Fig. 4 shows a schematic diagram of a UIA/GUS
7 expression cassette;
8 Fig. 5 shows a schematic diagram of a U2B"/GUS
9 expression cassette;
10 Fig. 6 shows a graph of fluorescence against time
11 plotted from data collected from a fluorometric
12 assessment of GUS gene expression controlled by
13 potato spliceosomal protein gene promoters;
14 Fig. 7 shows a schematic diagram of a UIA/GUS
15 expression cassette;
16 Fig. 8 shows a schematic diagram of U2B"/GUS
17 expression cassette;
18 Fig. 9 shows a schematic diagram of the genomic
19 organisation of the potato U1A gene; and
20 Fig. 10 shows a schematic diagram of the genomic
21 organisation of the potato U2B" gene.

22

23 Example 1

24 Genes for two plant spliceosomal proteins from the
25 dicotyledonous plant, potato (*Solanum tuberosum*) were
26 isolated encoding the spliceosomal proteins U1A and
27 U2B".

28

29 The promoter of a potato U2B" gene was isolated from a
30 potato genomic library in λ EMBL 3 by conventional
31 methods. The library was screened with a potato U2B"
32 cDNA clone (Simpson et al., 1991) using standard
33 procedures (Sambrook et al., 1989). The potato U2B"
34 genomic clone was plaque-purified, DNA prepared,
35 fragments subcloned into plasmid vectors and DNA
36 sequenced by standard procedures (Sambrook et al.

1 1989).

2

3 The genomic clone contained an insert of 15 kilobase
4 pairs (kbp) (Fig. 1) from which relevant sub-fragments
5 were cloned into plasmid vectors (Fig. 1). The clone
6 only contained a fragment of coding region of U2B" (100
7 bp), a fragment of the first intron in U2B" (100 bp)
8 and approximately 15 kbp of upstream sequences
9 containing the U2B" promoter.

10

11 Example 2

12 The promoter of a potato U1A gene was isolated from a
13 potato genomic library in λ EMBL 4. The library was
14 screened with a potato U2B" cDNA clone (Simpson et al.,
15 1991) because U1A and U2B" are closely related, and a
16 genomic clone was obtained by known methods. The
17 genomic clone was characterised by state-of-the-art
18 methodologies as described by Sambrook et al., (1989).
19 The clone contained an insert of 15 kbp (Fig. 2) from
20 which relevant sub-fragments were cloned into plasmid
21 vectors by standard techniques (Sambrook et al., 1989).
22 The clone contained the whole of the U1A coding
23 sequence on five exons, four introns and 7 and 2 kbp of
24 5' and 3' flanking sequence respectively.

25

26 The promoter regions can be linked to marker genes such
27 as bacterial β -glucuronidase (Jefferson, 1987) by
28 standard molecular techniques (Sambrook et al., 1989).
29 Promoter constructs can be analyzed by introduction
30 into plant cells by known methodology such as chemical
31 or electrical transfection, microinjection, biolistics
32 or *Agrobacterium*-mediated or other vector-mediated
33 transformation (see Shaw, 1988). Transgenic plants
34 containing the construct can be analyzed by detecting
35 the presence of GUS enzyme and thereby its expression
36 using known methods of histochemical staining

1 (Jefferson, 1987). Levels of expression of marker
2 genes driven by the promoters in either stably
3 transformed plants or transiently transformed plant
4 cells or protoplasts can be assessed by comparison with
5 levels of endogenous gene expression and of marker gene
6 expression driven by the Cauliflower Mosaic Virus 35S
7 RNA promoter. This analysis can use known, state-of-
8 the-art methodologies to detect RNA transcripts
9 (Sambrook et al., 1989; Simpson et al., 1992) and to
10 detect production of enzyme from marker genes, for
11 example, for GUS marker gene activity.

12
13 The promoter region can be incorporated into plasmid
14 vectors designed for general use in construct
15 production in *E.coli*, and for use in stable,
16 *Agrobacterium*-mediated transformation and in transient
17 transformation or stable, physical transformation
18 methods. DNA sequences to be expressed in the
19 transgenic plant can be inserted behind the promoter
20 regions as is currently commonly performed using the
21 CaMV 35S RNA promoter (see Shaw, 1988) prior to
22 introduction into plant cells or production of
23 transgenic plants.

24

25 Example 3

26 A gene for a plant spliceosomal protein was cloned from
27 the monocotyledonous plant, maize (*Zea mays* L.), which
28 encodes the spliceosomal protein PRP8 (Jackson et al.,
29 1988) by virtue of sequence homology to PRP8 of yeast.

30

31 The genomic clone of maize PRP8 was isolated from a
32 maize genomic library constructed in λ EMBL 4 by
33 conventional methods. The library was screened with a
34 fragment of maize PRP8 generated by state-of-the-art
35 methodologies, such as polymerase chain reaction (PCR)
36 amplification using oligonucleotide sequences designed

1 from the yeast and *Caenorhabditis elegans* PRP8 DNA
2 sequences, cloning and sequencing. The maize PRP8
3 genomic clone was plaque purified, DNA prepared and
4 fragments subcloned into plasmid vectors by standard
5 procedures.

6

7 The PRP8 promoter region can be incorporated into
8 plasmid vectors as previously described.

9

10 Example 4

11 **Promoter/GUS Constructions**

12 Two expression cassettes containing the β -glucuronidase
13 (GUS) gene driven by the potato U1A and U2B"
14 spliceosomal protein gene promoters respectively were
15 constructed. Both cassettes contained the nopaline
16 synthase poly A (NOS-ter) downstream to the GUS gene.
17 The construction of these expression cassettes was
18 achieved by replacing the CaMV 35S promoter in the
19 vector pBI221 by upstream sequences of the potato U1A
20 and U2B" spliceosomal protein genes. Genomic clones
21 described above were used.

22

23 **U1A**

24 A 4.5 kb *Pst*I fragment containing about 2.5 kb upstream
25 to the U1A gene, the first and the second exons, the
26 first intron and part of the second intron was used to
27 clone the U1A promoter. In order to clone the
28 promoter, a site specific mutation was generated which
29 changed the sequence ATGGCG at the translation start
30 site into TCTAGA to introduce an *Xba*I restriction site.
31 Subsequently, the entire 2.5 kb upstream region was
32 cloned into pBI221 after eliminating the 35S promoter
33 using the restriction enzymes *Pst*I and *Xba*I (fig. 4).

34

1 **U2B"**

2 An *EcoRI*/*SalI* fragment containing about 2 kb upstream
3 to the U2B" gene, the entire first exon, and part of
4 the first intron was used to clone the U2B" promoter.
5 Appropriate restriction sites were introduced into the
6 upstream region using PCR amplification. The first PCR
7 primer started a few nucleotides downstream of the
8 *EcoRI* site and contained additional nucleotides in its
9 5' end to provide *PstI* and *BglIII* sites. The second
10 primer started two nucleotides upstream of the ATG
11 start codon and contained additional nucleotides
12 providing a site for cleavage by *BamHI* at its 5' end.
13 After PCR amplification and digestion with *BamHI* and
14 *PstI*, the 2 kb upstream region was cloned into pBI221
15 replacing the 35S promoter (fig. 5).

16

17 **Protoplast Isolation**

18 The constructed cassettes were tested for promoter
19 activity in transient gene expression assays using
20 tobacco protoplasts. Protoplasts were prepared from
21 young, fully expanded tobacco leaves. Leaves were
22 placed in 7% Domestos for 10 minutes, washed with
23 sterile tap water, dried, and peeled to remove the
24 lower epidermis. Peeled leaf pieces were placed onto
25 15 ml enzyme solution in a sterile Petri dish [enzyme
26 solution: 1 mg/ml cellulase, 0.5 mg/ml driselase, 0.2
27 mg/ml macerase, suspended in TO^- (see appendix)] and
28 incubated overnight in the dark at 25°C. Protoplasts
29 were transferred into two sterile 10 ml tubes through
30 sterile sieves and spun at 400 rpm for 5 minutes.
31 After resuspending each protoplast pellet in 10 ml TO^- ,
32 5 ml aliquots were each layered onto 2.5 ml 16% sucrose
33 and spun at 800 rpm for 5 minutes. The purified
34 protoplasts were resuspended in 10 ml TO^- , spun at 400
35 rpm for 5 minutes, and recollected in 5 ml TO^- .

36

1 **Protoplast Transfection**

2 Transfection of the protoplasts with the U1A/GUS and
3 U2B"/GUS constructs was achieved using plasmid DNA
4 purified by Qiagen (Trade Mark) columns. For each
5 experiment 200 μ l of the purified protoplasts were
6 transfected with 30 μ g DNA dissolved in 20 μ l distilled
7 water. After dropwise addition of the DNA and careful
8 homogenization of protoplasts, 200 ml of PEG solution
9 was added. [PEG solution: 25% PEG 8000, 0.1M $\text{Ca}(\text{NO}_3)_2$,
10 0.45M Mannitol, 10 mM MES: 2-(N-Morpholino)
11 ethanesulfonic acid] was added dropwise. The
12 transfected protoplasts were then incubated at ambient
13 temperature for 20 minutes; subsequently, 4 ml calcium
14 nitrate was carefully added [0.275 M $\text{Ca}(\text{NO}_3)_2$]. After
15 an incubation period of 20 minutes at ambient
16 temperature protoplasts were spun at 400 rpm for 3
17 minutes, suspended, and resuspended in 5 ml TO^+ . The
18 transfected protoplasts were then incubated in sealed
19 Petri dishes in the light at 25°C for 48 hours.

20

21 **GUS Assays**

22 Samples were collected by spinning the protoplasts at
23 400 rpm for 3 minutes, resuspended, transferred into
24 Eppendorf tubes (1 ml each), and spun again at 1000 rpm
25 for 2 minutes in a cold microcentrifuge (4°C). The
26 pellet was then resuspended in 200 μ l extraction and
27 reaction buffer (50 mM NaPO_4 pH 7.0, 10 mM EDTA, 0.1%
28 Triton X-100, 10 mM β -mercaptoethanol), spun at 1300
29 rpm for 3-5 minutes, and the supernatant was stored in
30 clean Eppendorf tubes. 50 μ l sample extract was added
31 to 220 μ l extraction and reaction buffer in an
32 eppendorf tube and preincubated at 37°C for 5 minutes.
33 At 15 sec intervals, 30 μ l 1 mM MUG (4-methyl
34 umbelliferyl glucuronide) substrate was added to each
35 tube giving time to stop reactions at accurate
36 intervals. At defined time points, 100 μ l samples were

1 taken and added to 2.9 ml stop solution (0.2 M Na_2CO_3)
2 in a clean cuvette. At this stage, GUS activity could
3 be inspected visually by transillumination using a long
4 wave UV light box. Quantitative measurements of
5 fluorescence were made using a fluorometer.

6
7 Protoplast transfections and GUS assays have been
8 performed to investigate the expression of the GUS gene
9 driven by splicesomal protein gene promoters using the
10 above described constructs. In all experiments, the
11 vector pBI221 which contains the GUS gene driven by the
12 CaMV 35S promoter and terminated by the nopaline
13 synthase poly A was used as a positive control.
14 Negative controls were a: extracts from protoplasts
15 transfected with water instead of plasmid DNA; and b:
16 MUG substrate added to the extraction and reaction
17 buffer without the addition of any extracts. Reactions
18 were carried out for 30 minutes, 1 hour, 2 hours, 3
19 hours, 4 hours, 5 hours and 20 hours. In order to
20 assess the activity of the tested promoters in
21 comparison to the 35S promoter, fluorometer readings
22 were set at 100 for pBI221 at each time point so that
23 the readings obtained for U1A and U2B" indicated the
24 GUS activity in each case relative to that obtained for
25 the 35S promoter. The results which are illustrated in
26 fig 6 show that the U2B" promoter has about 70% of the
27 activity of the 35S promoter whereas the U1A promoter
28 gave rise to 10-15% of the activity observed for the
29 35S promoter.

30
31 Figs. 9 and 10 show the genomic organisation of the
32 original genomic clones made of potato genes U1A and
33 U2B". The promoter of the U1A gene is believed to lie
34 in the region between the ATG initiation codon and the
35 PstI site 2.5 kb upstream of the coding sequence
36 initiating at the ATG start codon. The promoter of the

1 U2B" gene is believed to lie in the region between the
2 ATG coding region and the *EcoRI* site 2.0 kb upstream of
3 the coding sequence initiated by the ATG start codon.
4

5 Figs. 9 and 10 indicate regions of the genomic clones
6 containing the promoters of U1A and U2B" which have
7 been sequenced. The sequences of those regions are
8 shown in the sequence listing.
9

10 The partial U1A promoter sequence is shown on the
11 sequence listing as SEQ:ID:No1. Sequence data is
12 presented for the 3' end of the region between the ATG
13 codon and a *PstI* site 2.5kb upstream. About 1.7kb of
14 the 5' end of this region has not been sequenced. The
15 partial sequence of the U2B" genomic clone is shown in
16 the sequence listing as three separate sequences since
17 incomplete sequence data is available for this
18 promoter. The first of the U2B" sequences shown is
19 SEQ:ID:No2 which corresponds to the region from the 5'
20 end TCT to a TAA 119 bases downstream. After the 3'
21 end of SEQ:ID:No2 there is a region of approximately
22 900 bases which have not been sequenced. After the 3'
23 end of the 900bp unidentified region, the next U2B"
24 sequence portion available is SEQ:ID:No3 which runs
25 from a TCT approximately 900 bases downstream of the
26 end of SEQ:ID:No2 to an ATC a further 139 bases
27 downstream. After the 3' end of SEQ:ID:No3 there is a
28 region of 11 bases which have not been identified.
29 This unidentified region is followed by the last
30 sequenced portion of the U2B" clone which is designated
31 as SEQ:ID:No4 which starts at ATA at its 5' end and
32 ends at the ATG codon (bases 729-731) initiating
33 transcription of the coding region of U2B" (not shown).
34

35 Spliceosomal protein gene promoters (or other
36 expression control polynucleotides) have the advantages

1 that (a) spliceosomal proteins are absolutely required
2 and thus spliceosomal protein gene promoters are likely
3 to be active in every cell and tissue type; (b) they
4 are not derived from infectious agents which overcomes
5 objections to the use of such sequences due to
6 potential recombination; and (c) different genes or DNA
7 sequences are likely to be expressed at different
8 levels reflecting the relative abundance of the
9 different spliceosomal proteins.

10

11 Modifications and improvements may incorporated without
12 departing from the scope of the invention. For
13 example, an expression control polynucleotide in
14 accordance with the invention may be operably linked to
15 a second polynucleotide which has some function other
16 than coding for a polypeptide. One such example might
17 be to operably link an expression control
18 polynucleotide to a gene encoding a ribozyme or anti-
19 sense RNA. Indeed it will be realised by the skilled
20 man that the function of the gene under the control of
21 the expression control polynucleotide of the invention
22 is not important, and that the expression of many
23 diverse genes can be controlled.

1 References

2 The following references are incorporated herein by
3 reference.

4
5 Jackson, S.P., Lossky, M & Beggs, J.D. (1988). Mol
6 Cell Biol. 8, 1067-1075.

7
8 Jefferson, R.A. (1987). Assaying chimaeric genes in
9 plants: the GUS gene fusion system. Plant Molecular
10 Biology Reporter 5, 387-405.

11
12 Simpson, G.G., Vaux, P., Clark, G., Waugh, R., Beggs,
13 J.D. & Brown, J.W.S. (1991). Evolutionary conservation
14 of the spliceosomal protein, U2B". Nucleic Acids
15 Research 19, 5213-5217.

16
17 Simpson, C.G., Sinibaldi, R. & Brown, J.W.S. (1992).
18 Rapid analysis of plant gene expression by a novel
19 reverse transcriptase-PCR method. Plant Journal 2,
20 835-836.

21
22 Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989).
23 Molecular Cloning - A Laboratory Manual. Cold Spring
24 Harbor Laboratory Press, Cold Spring Harbor, New York.

25
26 Shaw, C.H. (1988). Plant molecular biology: a
27 practical approach. IRL Press.

28

1 Appendix 1

2 TO-:

3 4 ml solution 1

4 200 ml solution 3

5 200 ml solution 4

6 200 ml NAA (3 mg/ml)

7 200 ml BAP (1 mg/ml)

8 16 g Mannitol

9 pH 5.5 adjusted with NaOH

10 filtersterilised +200 ml

11 cefotaxime (100 mg/ml)

12

13

14

15

16

17 **Solution 1:**18 10.30 mM NH_4NO_3 19 9.40 mM KNO_3 20 1.50 mM $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 21 0.75 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 22 0.62 mM KH_2PO_4

23

24 **Solution 3:**25 16.00 μM H_3BO_3 26 0.60 μM $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 27 3.50 μM $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 28 0.12 μM $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 29 0.22 μM AlCl_3 30 0.13 μM $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ 31 0.06 μM KI

32

TO+:

4 ml solution 1

4 ml solution 2

200 ml solution 3

200 ml solution 4

200 ml NAA (3 mg/ml)

200 ml BAP (1mg/ml)

16 g Mannitol

4 g Sucrose

40 ml Tween 20

pH 5.5 adjusted with NaOH

filtersterilised +200 ml

cefotaxime (100 mg/ml)

Solution 2:100 μM FeSO_4 100 μM Na_2 EDTA**Solution 4:**555.00 μM Inositol3.00 μM Thiamine5.00 μM Pyridoxine8.00 μM Niacin2.00 μM Pantothenate0.04 μM Biotine+ 1 mg/ μl NaOH

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Scottish Crop Research Institute
- (B) STREET: Invergowrie
- (C) CITY: Dundee
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): DD2 5DA
- (G) TELEPHONE: +44 1382 562731
- (H) TELEFAX: +44 1382 561442

(ii) TITLE OF INVENTION: Expression control polynucleotides derived from spliceosomal protein gene promoters

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 784 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Solanum tuberosum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TTAGAATTAG AATCCCCATT TTTAAGAATA ATCCTAGATA ATTTCTTAA ACATGACAAT	60
TGATACCCAC AATTAATTAC TATTACATAA ATTTTACCT AAATTAGATA TAACTTTCAA	120
TTTCAAAAAT TAAACCCAA AAAAATTGAA CGACAATACG AGAGGGGATC AAACATAGGC	180

20

GAGCAATTAG AGAAATTGAC GGGTAGACAT CAACAAACCA TCAAGAATTT AAAAGCGGAA	240
AGAGAAAAAA ATACACTATG GACGAATATT TTTATAGAAT TCAATATGTA AACTAATAA	300
ACAAGAAAGT AAATCATCTT TTATTCAAAG TAATGAAGAA GAAGAATTGA ATAAATATTT	360
ACATAATCAA TAAAAAAAC TCATTCAAAA GAATCGTGTG TATGGGAAAG AAGAAGAAAA	420
AAAAGGCAGA AAAAAACCAC TTCCCAATAA AAAAGGACAT CATGCTGCCA CCTCCTAAAA	480
TTATTTAATT TAATTAAAA AAAAACTTCC CAACACGTGG GCTACTAATT GCAAAATTTA	540
ATTTTAAAA AGCTTTTTTT GTCAAGAAAA TAAAAGATGG CTATATGTTG CCAATTAGTA	600
AAATGGGATG TCATGCTGTG TCATTTTTTC TTGAGTTGTT AAGGGCTCAA AGCCCAATTG	660
TTTATCCAGC CCAAGCCCAA ATCGGAGCCC TATTCGTGCC CAAAAATTTT TGGAGAAATT	720
AACGACAACT GAAGTTTCTA CCTCACCGGC GAAAGTTGCA GCTAGGCGTA GACGAGGAGC	780
CATG	784

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Solanum tuberosum*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TCTCTCCAGC TCTTCCCTCC TAAAACAACC ATTTTATGAG TACAGACACA AACCAGCTTA	60
GCAACCAGTA AATCCAAAAC TTTAATTCCA CGTGTAAGCG CTAACACTTC ACCCACTAA	119

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 139 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Solanum tuberosum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TCTTTACACT ATTTTCAAAT AACGATGAGA CCTGTAATAA TGTAATAACT TGAAAATAGA	60
ACAATAACTC ATTCAGTACA ACAAATAAAA TCATACTAAT GTATATTTTT AAAACAATT	120
TAATCTATT TAATTAATC	139

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 731 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Solanum tuberosum

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATATCTAAAA AAATGTGATT GGAGCACTCA TTGACCACAC GAATGAATCT ACAATGTAGA	60
TCATTCTATC TAGACAAATA ATGCAAAAAC TAAAAGATAA AAGTAATATT ATACTACAAT	120
CTGTTAAATG ATATCAATAT TACAAAAGTT CTCTACGATG TCAATACATA TATTAAAATC	180
TATTTGATTA ATCAGAAACA TATCATGTGT GAATTTTTTT AATTAAAGAT CCCTTTAATC	240
ATCTGAATCA ACCTTGGCTG GTCTCACATC TTTCCACCCT CTACTCGGTC TTCACTTTCT	300

CTTAACTAG GGAAGAACA CATGATATTA GCTTAGATTA ATTAAACAAG CTCATCAAAA	360
CTACCATCCA ATTTAAGCCA ATAATGTTA AATAAAACA AAAACAACGT ACTCATTTTT	420
TCATAACAAG AAGTTAAAT TTATATGAAT CCTTACTCCA AAAAGAAGA AAAATTACAA	480
TATCAATATA TATACTACT CTATTTGGTT AGTCAACAAA ATGTTAGTAT ATGTATTGCA	540
AGTTCGCAAC ACCCGCTTGG GCCTTGACCA CATATTTATA TGGGCCGGTT GTCAATTTAA	600
GCCCACTTTG TGTTGTTTCG CTTTCTTGT AGCTCCAAAC TCTTGAAAT TTGTCGAGCA	660
CATTCAGAAA TCACAGAGAA GAGCAAGTGA ATATACATAC AGATAGAGAA AAGCTGCTCT	720
GCTCGGTAAT G	731

1 Claims

2

3 1 An expression control polynucleotide at least
4 partially derived from a spliceosomal protein gene
5 promoter.

6

7 2 An expression control polynucleotide according to
8 Claim 1 derived from a spliceosomal protein gene
9 promoter of a plant.

10

11 3 A recombinant polynucleotide comprising an
12 expression control polynucleotide according to
13 Claim 1 or Claim 2 operably linked to a second
14 polynucleotide.

15

16 4 A recombinant polynucleotide according to Claim 3,
17 wherein the second polynucleotide encodes a
18 polypeptide.

19

20 5 A recombinant polynucleotide according to Claim 3,
21 wherein the second polynucleotide encodes a
22 ribozyme.

23

24 6 A recombinant polynucleotide according to Claim 3,
25 wherein the second polynucleotide encodes anti-
26 sense RNA.

27

28 7 A recombinant vector containing an expression
29 control polynucleotide according to Claim 1 or
30 Claim 2, or containing a recombinant
31 polynucleotide according to any one of Claims 3 to
32 6.

33

34 8 A method for producing a recombinant vector, said
35 method comprising ligating an expression control
36 polynucleotide according to Claim 1 or Claim 2 to

- 1 a vector.
- 2
- 3 9 A transformed host cell containing a recombinant
- 4 polynucleotide according to any one of Claims 3 to
- 5 6, or a recombinant vector as claimed in Claim 7.
- 6
- 7 10 A transgenic organism and/or the progeny and/or
- 8 seeds thereof, containing a recombinant
- 9 polynucleotide according to any one of Claims 3 to
- 10 6, or a recombinant vector according to Claim 7.
- 11
- 12 11 A method for controlling the expression of a
- 13 polypeptide from a polynucleotide encoding the
- 14 polypeptide, said method comprising operably
- 15 linking said polynucleotide to an expression
- 16 control polynucleotide according to Claim 1 or
- 17 Claim 2.
- 18
- 19 12 A spliceosomal protein gene promoter deposited as
- 20 any one of NCTC 12864, NCTC 12865 and NCTC 12866.
- 21
- 22 13 A recombinant polynucleotide having a spliceosomal
- 23 protein gene promoter according to Claim 12.
- 24
- 25 14 A recombinant vector having a spliceosomal protein
- 26 gene promoter according to Claim 12.
- 27
- 28 15 A transformed host cell containing a spliceosomal
- 29 protein gene promoter as claimed in Claim 12, a
- 30 recombinant polynucleotide as claimed in Claim 13,
- 31 and/or a recombinant vector as claimed in Claim
- 32 14.
- 33
- 34 16 A spliceosomal protein gene promoter comprising a
- 35 region of a plasmid in NCTC 12864 between an ATG
- 36 codon at the start of the coding sequence and a

- 1 PstI site 2.5 kb upstream of the ATG codon.
2
3 17 A spliceosomal protein gene promoter comprising a
4 region of a plasmid in NCTC 12865 between an ATG
5 codon at the start of the coding sequence and an
6 EcoRI site 2 kb upstream of the ATG codon.
7
8 18 A polynucleotide having any one of the sequences
9 SEQ:ID:No1, SEQ:ID:No2, SEQ:ID:No3, SEQ:ID:No4 and
10 functional equivalents thereof.

1/7

FIG. 1

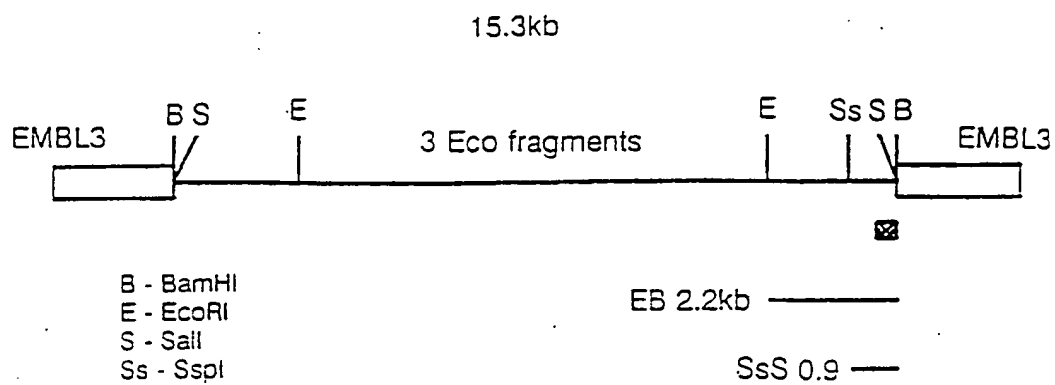
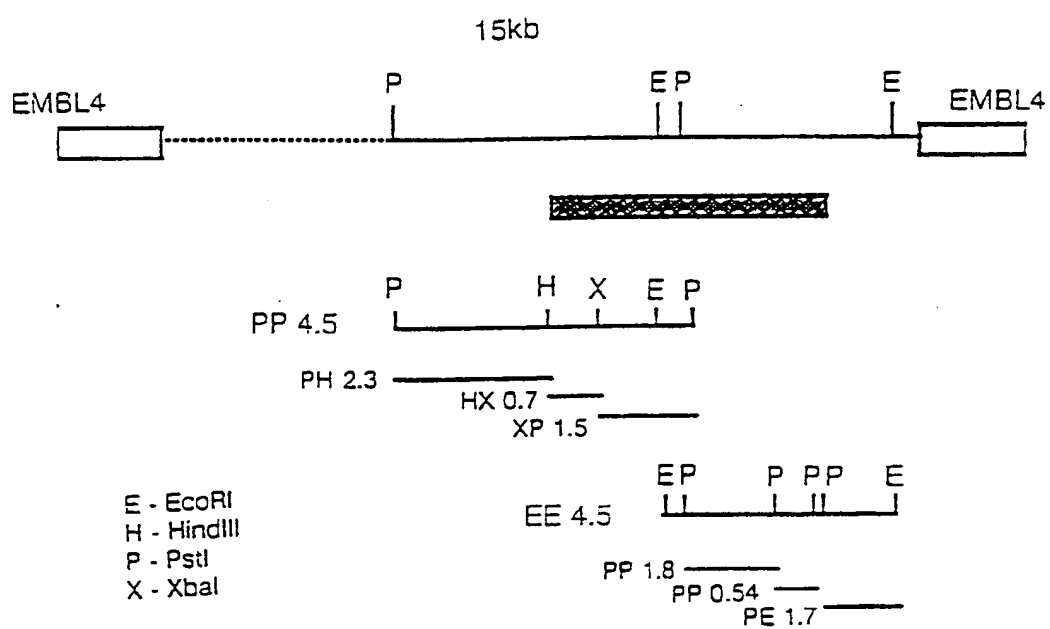
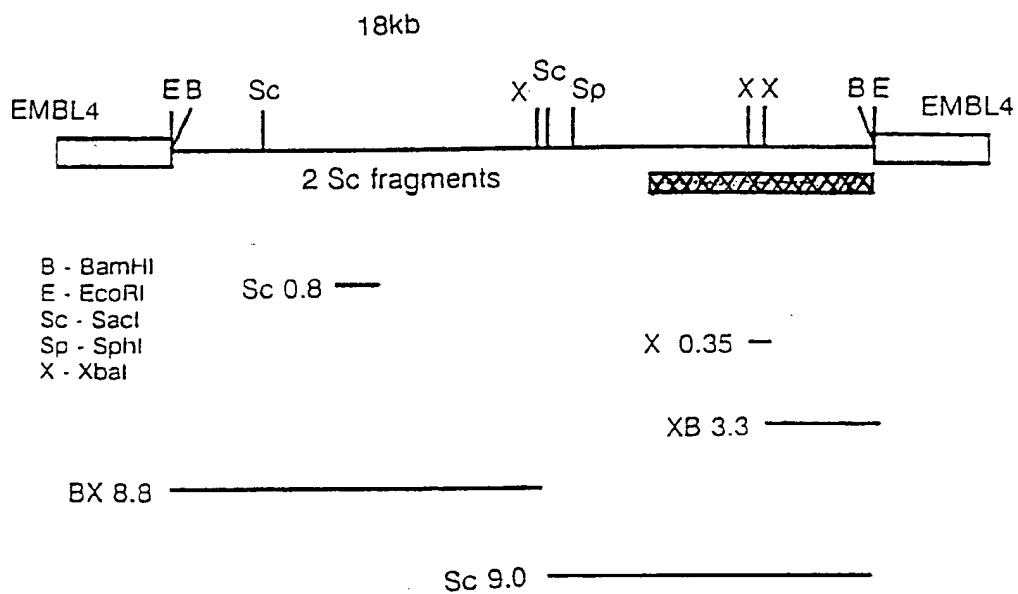


FIG. 2



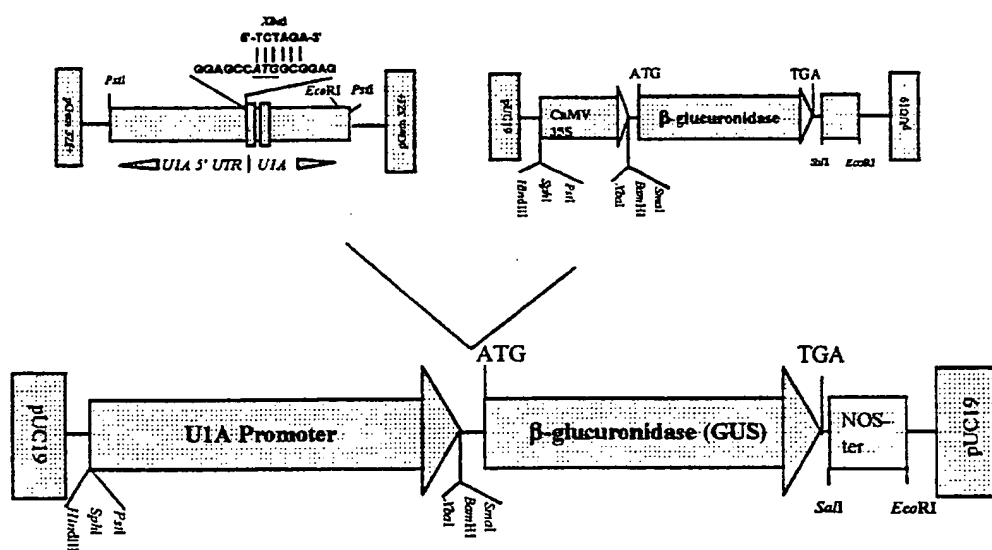
2/7

FIG. 3



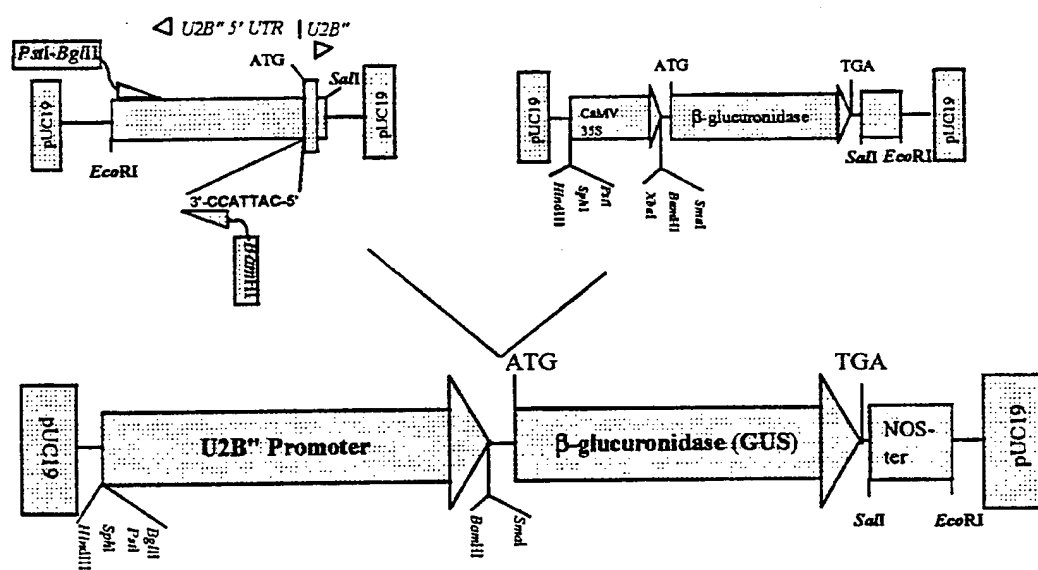
3/7

FIG. 4



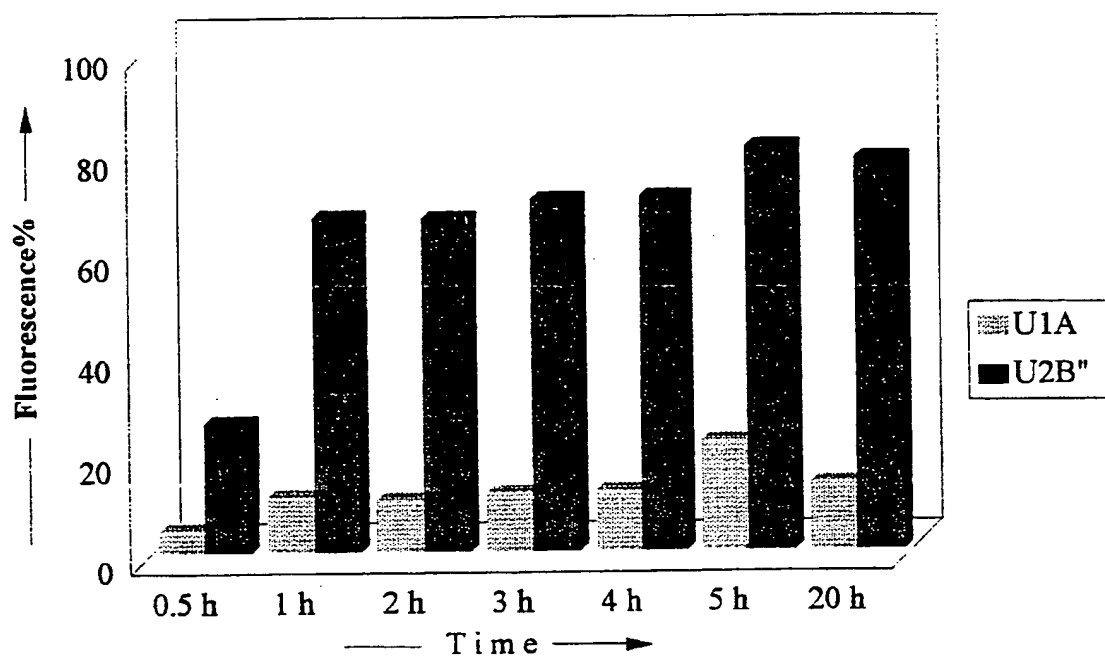
4/7

FIG. 5



5/7

FIG. 6



6/7

FIG. 7

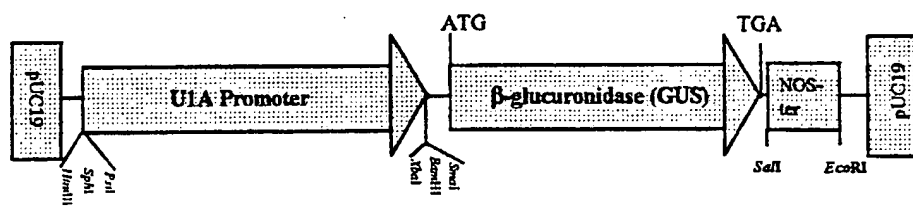
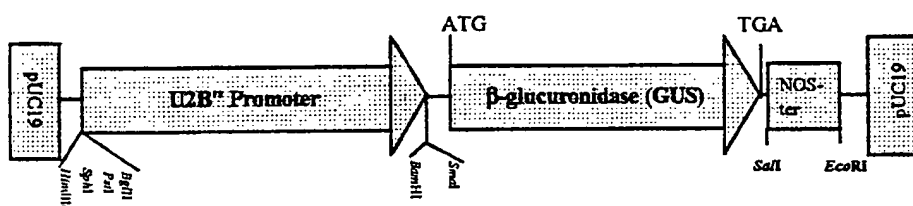


FIG. 8



7/7

FIG. 9

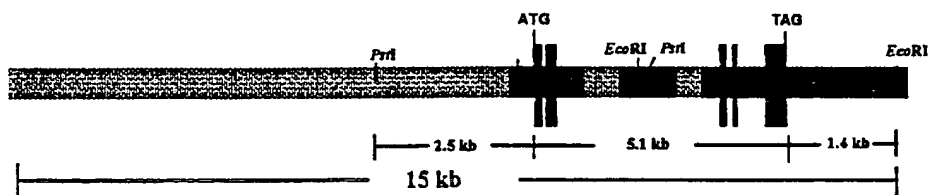
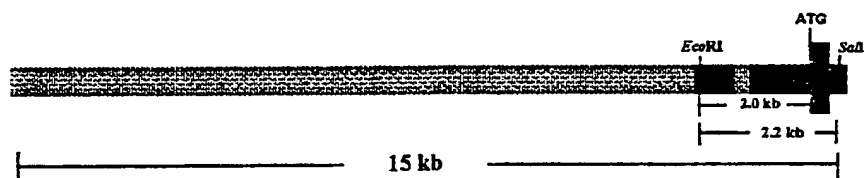


FIG. 10



■ Sequenced
▨ Unsequenced

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 95/01443

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N15/29 A01H5/00 C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MOL. CELL. BIOL., vol. 8, no. 3, 1988 pages 1067-1075, JACKSON, S.P., ET AL. 'CLONING OF THE RNA8 GENE OF SACCHAROMYCES CERVISIAE, DETECTION OF THE RNA8 PROTEIN, AND DEMONSTRATION THAT IT IS ESSENTIAL FOR NUCLEAR PRE-MRNA SPLICING' see page 1068, right column, last paragraph ---	1
X	NUCLEIC ACIDS RESEARCH, vol. 22, no. 11, 11 June 1994 pages 1996-2002, BIAMONTI, G., ET AL. 'TWO HOMOLOGOUS GENES, ORIGINATED BY DUPLICATION, ENCODE THE HUMAN HNRNP PROTEINS A2 AND A1' see page 2001, left column, paragraph 1 ---	1,3,4, 7-11
-/--		

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

* Special categories of cited documents :

- 'A' document defining the general state of the art which is not considered to be of particular relevance
- 'E' earlier document but published on or after the international filing date
- 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- 'O' document referring to an oral disclosure, use, exhibition or other means
- 'P' document published prior to the international filing date but later than the priority date claimed

- 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- '&' document member of the same patent family

Date of the actual completion of the international search

2 November 1995

Date of mailing of the international search report

22. 11. 95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+ 31-70) 340-3016

Authorized officer

Maddox, A

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 95/01443

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J. MOL. BIOL., vol. 230, 5 March 1993 pages 77-89, BIAMONTI, G., ET AL. 'HUMAN HNRNP PROTEIN A1 GENE EXPRESSION STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF THE PROMOTER' see the whole document ---	1,3,4, 7-11
P,X	JOURNAL OF EXPERIMENTAL BOTANY, vol. 46, no. SUPPL, 1995 page 38 HAMILTON, J.I., ET AL. 'CHARACTERISATION OF A GENE ENCODING THE SPLICEOSOMAL PROTEIN PRP8 FROM MAIZE' see abstract & ANNUAL MEETING OF THE SOCIETY FOR EXPERIMENTAL BIOLOGY, ST. ANDREWS, SCOTLAND, UK, APRIL 3-7, 1995., ---	1,2
A	PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY OF LONDON B BIOLOGICAL SCIENCES 342 (1301). 1993. 217-224., BROWN J W S , ET AL. ' Plant pre-mRNA splicing and splicing components.' see page 219, left column see page 222, right column ---	1-18
A	NUCLEIC ACIDS RESEARCH, vol. 19, no. 19, 1991 pages 5213-5217, SIMPSON, G.G., ET AL. 'EVOLUTIONARY CONSERVATION OF THE SPLICEOSOMAL PROTEIN, U2B'' cited in the application see the whole document ---	1-18
A	FEBS LETTERS, vol. 318, no. 1, February 1993 pages 4-6, KULESZA, H., ET AL. 'DETECTION OF A PLANT PROTEIN ANALAGOUS TO THE YEAST SPLICEOSOMAL PROTEIN, PRP8' see the whole document ---	1-18
A	NUCLEIC ACIDS RESEARCH, vol. 18, no. 8, 1990 pages 1941-1949, KISS, T., ET AL. 'MOLECULAR ANALYSIS OF A U3 RNA GENE LOCUS IN TOMATO: TRANSCRIPTION SIGNALS, THE CODING REGION, EXPRESSION IN TRANSGENIC TOBACCO PLANTS AND TANDEMLY REPEATED PSEUDOGENES' see page 1946, right column ---	1-18

-/--

INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 95/01443

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>NUCLEIC ACIDS RESEARCH, vol. 21, no. 7, 1993 pages 1533-1540, EDOH, D., ET AL. 'ACTIVITY OF U-SNRNA GENES WITH MODIFIED PLACEMENT OF PROMOTER ELEMENTS IN TRANSFECTED PROTOPLASTS AND STABLY TRANSFORMED TOBACCO' see the whole document -----</p>	1-18

